

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application:

1999年12月15日

出 願 番 号
Application Number:

平成11年特許願第355382号

出 願 人
Applicant (s):

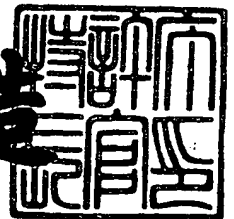
日本電気株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年10月27日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2000-3088638

【書類名】 特許願

【整理番号】 49210395

【提出日】 平成11年12月15日

【あて先】 特許庁長官殿

【国際特許分類】 H04L 12/28

【発明者】

 【住所又は居所】 東京都港区芝五丁目7番1号 日本電気株式会社内

 【氏名】 神谷 聡史

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100088812

 【弁理士】

 【氏名又は名称】 ▲柳▼川 信

【手数料の表示】

 【予納台帳番号】 030982

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 パケットスイッチ及びパケットスイッチング方法

【特許請求の範囲】

【請求項 1】 ネットワークにおけるデータのフローを制御するスイッチであって、

複数の入力ポートと、

複数の出力ポートと、

前記複数の出力ポートのうちの指定出力ポートへデータを送るように、前記複数の入力ポートのうちの特定の入力ポートを予約し、入力ポートと出力ポートとの接続状態を決定する N 個（N は自然数、以下同じ）の入力ポートスケジューリングモジュールを有するスケジューラと、

を含み、

前記スケジューラは、

タイムスロット単位に、

各入力ポートスケジューリングモジュールが、前段のスケジューリングモジュールから、ある予約タイムスロットの予約状況情報を受信し、

各入力ポートスケジューリングモジュールが、その予約タイムスロットにおける当該入力ポートスケジューリングモジュールからのパケット送出予約可否を決定し、

各入力ポートスケジューリングモジュールが、前段のスケジューリングモジュールから受信した予約状況情報に、自スケジューリングモジュールの予約結果を反映させて、次段のスケジューリングモジュールに送信する手段と、

N 個のタイムスロットを単位とするフレームを定義して、前記フレーム時間内で、前記フレームの次のフレーム中の N 個のタイムスロットでの予約を行う手段と、

現在のスケジューリングモジュールが、前のスケジューリングモジュールから予約状況情報を受信する手段と、

前記現在のスケジューリングモジュールが、前記複数の出力ポートのうちの 1

つにアクセスしようとする将来のタイムスロットを予め次のフレーム内の特定のタイムスロットに決定する手段と、

前記複数の出力ポートのうちの 1 つを前記将来のタイムスロットでの送信用に予約するように選択する手段と、

前記将来のタイムスロットが他のスケジューリングモジュールによって予約済みかどうかを判断する手段と、

前記将来のタイムスロットが他のスケジューリングモジュールによって予約されていない場合には、前記将来のタイムスロットを予約し、前記将来のタイムスロットが予約されたことを示す情報を前記予約状況情報に入れる手段と、

前記予約状況情報を次のスケジューリングモジュールに渡す手段と、

タイムスロットでの予約過程の観点から見た場合、

前記予約過程が、フレームの先頭で同時に開始され、フレーム内で同時にパイプライン処理により進行し、

フレームの末端で同時に完了し、

前記入力ポートスケジューリングモジュールが、フレームの先頭で同時に開始する複数の前記予約過程それぞれにおいて、次のフレーム内の各々異なる予約タイムスロットを対象として処理を開始する手段と、

前記複数のスケジューリングモジュールによる前記予約の順序を変更する予約順序変更手段とを含み、

この変更された順序で前記複数のスケジューリングモジュールそれぞれが各フレーム毎に次のフレームの packets について出力すべきポートを予約するようにしたことを特徴とする packets スイッチ。

【請求項 2】 前記複数のスケジューリングモジュールは他のモジュールとの論理的接続順序に対応した順序で前記予約を実行するものであり、前記予約順序変更手段は前記複数のスケジューリングモジュールの接続トポロジを変更することを特徴とする請求項 1 記載の packets スイッチ。

【請求項 3】 前記予約順序変更手段は、前記複数のスケジューリングモジュールの論理的接続状態を変更するようにスイッチング動作するスイッチと、このスイッチのスイッチング動作を制御する制御データを記憶したテーブルとを含

むことを特徴とする請求項 2 記載の PACKET SWITCH。

【請求項 4】 前記複数のスケジューリングモジュールと前記スイッチ間の物理的接続は、電氣的接続であることを特徴とする請求項 3 記載の PACKET SWITCH。

【請求項 5】 前記複数のスケジューリングモジュールと前記スイッチ間の物理的接続は、光学的接続であることを特徴とする請求項 3 記載の PACKET SWITCH。

【請求項 6】 前記テーブルは、前記複数のスケジューリングモジュールそれぞれの内部に設けられていることを特徴とする請求項 3 ～ 5 のいずれかに記載の PACKET SWITCH。

【請求項 7】 前記テーブルは、前記複数のスケジューリングモジュールに対して共通に設けられていることを特徴とする請求項 3 ～ 5 のいずれかに記載の PACKET SWITCH。

【請求項 8】 前記制御データは、前記複数のスケジューリングモジュール各々がフレーム内の先頭において予約を開始するタイムスロットを各フレーム毎に変更できるように前記スイッチのスイッチング動作を制御するデータであることを特徴とする請求項 3 ～ 7 のいずれかに記載の PACKET SWITCH。

【請求項 9】 前記制御データは、前記予約を開始する予約開始スロットの前記複数のスケジューリングモジュールによる使用頻度を均等とするスケジューリングを実現するデータであることを特徴とする請求項 3 ～ 7 のいずれかに記載の PACKET SWITCH。

【請求項 10】 前記制御データは、前記予約を開始する予約開始スロットの前記複数のスケジューリングモジュールによる使用順序及び使用頻度を変更可能とするスケジューリングを実現するデータであることを特徴とする請求項 3 ～ 7 のいずれかに記載の PACKET SWITCH。

【請求項 11】 N 個の入力ポートスケジューリングモジュールを有するスイッチのスケジューラにて使用する、複数の出力ポートのうちの指定出力ポートへデータを送るように、複数の入力ポートのうちの特定の入力ポートを予約し、入力ポートと出力ポートとの接続状態を決定する PACKET SWITCHING 方法であ

って、

各入力スケジューリングモジュールが、前段のスケジューリングモジュールからあるタイムスロットの予約状況情報を受信するステップと、

各スケジューリングモジュールが、その予約タイムスロットにおける当該入力ポートスケジューリングモジュールからのパケット送出予約可否を決定するステップと、

各入力スケジューリングモジュールが、前段のスケジューリングモジュールから受信した予約状況情報に、自スケジューリングモジュールの予約結果を反映させて、次段のスケジューリングモジュールに送信するステップと、

を含み、

前記スケジューラは、

現在のスケジューリングモジュールが、前のスケジューリングモジュールから予約状況情報を受信するステップと、

前記現在のスケジューリングモジュールが、前記複数の出力ポートのうちの1つにアクセスしようとする将来のタイムスロットを予め次のフレーム内のタイムスロットに決定するステップと、

前記複数の出力ポートのうちの1つを前記将来のタイムスロットでの送信用に予約するように選択するステップと、

前記将来のタイムスロットが他のスケジューリングモジュールによって予約済みかどうかを判断するステップと、

前記将来のタイムスロットが他のスケジューリングモジュールによって予約されていない場合には、前記将来のタイムスロットを予約し、前記将来のタイムスロットが予約されたことを示す情報を前記予約状況情報に入れるステップと、

前記予約状況情報を次のスケジューリングモジュールに渡すステップと、

を実行し、

タイムスロットでの予約過程の観点から見た場合、

前記予約過程が、フレームの先頭で同時に開始され、フレーム内で同時にパイプライン処理により進行し、

フレームの末端で同時に完了し、

前記入力ポートスケジューリングモジュールが、フレームの先頭で同時に開始する複数の前記予約過程それぞれにおいて、次のフレーム内の各々異なる予約タイムスロットを対象として処理を開始し、

さらに、前記複数のスケジューリングモジュールによる前記予約の順序を変更し、この変更された順序で前記複数のスケジューリングモジュールそれぞれが各フレーム毎に次のフレームの packets について出力すべきポートを予約するようにしたことを特徴とする packets スイッチング方法。

【請求項 1 2】 前記複数のスケジューリングモジュールは他のモジュールとの論理的接続順序に対応した順序で前記予約を実行するものであり、前記予約順序の変更においては前記複数のスケジューリングモジュールの接続トポロジを変更することを特徴とする請求項 1 1 記載の packets スイッチング方法。

【請求項 1 3】 前記予約の順序の変更においては、前記複数のスケジューリングモジュールの論理的接続状態を変更するようにスイッチング動作するスイッチと、このスイッチのスイッチング動作を制御する制御データを記憶したテーブルとを用いて予約順序を変更することを特徴とする請求項 1 1 記載の packets スイッチング方法。

【請求項 1 4】 前記複数のスケジューリングモジュールと前記スイッチ間の物理的接続は、電氣的接続であることを特徴とする請求項 1 3 記載の packets スイッチング方法。

【請求項 1 5】 前記複数のスケジューリングモジュールと前記スイッチ間の物理的接続は、光学的接続であることを特徴とする請求項 1 3 記載の packets スイッチング方法。

【請求項 1 6】 前記テーブルは、前記複数のスケジューリングモジュールそれぞれの内部に設けられていることを特徴とする請求項 1 3 ～ 1 5 のいずれかに記載の packets スイッチング方法。

【請求項 1 7】 前記テーブルは、前記複数のスケジューリングモジュールに対して共通に設けられていることを特徴とする請求項 1 3 ～ 1 5 のいずれかに記載の packets スイッチング方法。

【請求項 1 8】 前記制御データは、前記複数のスケジューリングモジュール

ル各々がフレーム内の先頭において予約を開始するタイムスロットを各フレーム毎に変更できるように前記スイッチのスイッチング動作を制御するデータであることを特徴とする請求項 1 1 ~ 1 3 のいずれかに記載のパケットスイッチング方法。

【請求項 1 9】 前記制御データは、前記予約を開始する予約開始スロットの前記複数のスケジューリングモジュールによる使用頻度を均等とするスケジューリングを実現するデータであることを特徴とする請求項 1 3 ~ 1 8 のいずれかに記載のパケットスイッチング方法。

【請求項 2 0】 前記制御データは、前記予約を開始する予約開始スロットの前記複数のスケジューリングモジュールによる使用順序及び使用頻度を変更可能とするスケジューリングを実現するデータであることを特徴とする請求項 1 3 ~ 1 8 のいずれかに記載のパケットスイッチング方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明はパケットスイッチに関し、特にパケット交換システムにおける入出力間のスケジューリング及びそのスケジューリングを実現するためのパケットスイッチに関する。

【0 0 0 2】

【従来の技術】

近年のパケット交換システムでは、N 入力、N 出力（N は自然数、以下同じ）を有し、かつ各入力部が N 個の仮想出力キュー（Virtual Output Queuing：VOQ）を有する入力バッファ型スイッチが用いられることがある。

【0 0 0 3】

図 1 1 に一般的な N 入力、N 出力（N は自然数、以下同じ）の入力バッファ型パケットスイッチの構成が示されている。同図を参照すると、パケットスイッチ 4 0 は、データを入力する複数の入力ポートと、データを出力する複数の出力ポートと、入力ポートから入力されたデータをスイッチングして出力ポートへ転送

するデータスイッチング素子 5 4 と、スイッチ素子 5 4 を制御するスケジューラ 5 0 とを備えている。

【0 0 0 4】

入力ポートは仮想出力キュー (VOQ) 5 2 の構成をとっている。スイッチング素子 5 4 としてはクロスバースイッチが考えられる。スケジューラ 5 0 は分散スケジューリングの構成をとっており、入力ポート毎の分散スケジューリングモジュール 5 1 - i ($i = 1 \sim N$) から構成される。

【0 0 0 5】

上述したパケットスイッチは、クロスバースイッチ内の転送を、固定サイズのパケットにて行っているとする。これにより、スイッチシステムの動作時刻を量子化している。この量子化単位をタイムスロットと呼ぶ。

【0 0 0 6】

スケジューラは、タイムスロット単位に各入力ポートから各出力ポート毎の接続要求情報 (REQ) を受信し、接続要求情報に基づいて入力ポートと出力ポートの間の接続許可情報 (GRANT) を決定する。スケジューラは接続許可情報を基に入力ポートと出力ポートとの接続情報 (MSEL) を生成してスイッチング素子に通知し、スイッチング素子の入出力の接続を設定する。

【0 0 0 7】

また、スケジューラは接続許可情報を基に各入力ポートがどの出力ポートからのデータ転送が許可されているかを示す転送許可情報 (DSTMSG) を作成し、各入力ポートに対して転送許可情報を通知する。入力ポートは転送許可情報に従ってデータをスイッチング素子へ出力し、出力ポートがデータを受信してスイッチングが完了する。

【0 0 0 8】

スケジューラの目的は、 $N \times N$ の接続要求情報から $N \times N$ の接続許可情報を生成することである。接続許可情報を生成するに当たり、各分散スケジューリングモジュール 5 1 - 1 ~ 5 1 - N は、個々の入力ポートに対する出力ポートへの接続可否を決定している。

【0 0 0 9】

ある分散スケジューリングモジュール 51-n (n は自然数で、 $1 \leq n \leq N$) が接続許可とした出力ポートは、他の分散スケジューリングモジュール 51-m ($m \neq n$) にとっては別の分散スケジューリングモジュールに「予約」されたポートであり、接続許可を発出することが不可能なポートとなる。以下、ある分散スケジューリングモジュールがある出力ポートへ接続許可と決定する動作を「出力ポートを予約する」と表現する。

【0010】

パケットスイッチの分散型スケジューリングアルゴリズムとして、1999年 Globecom において A. Smiljanic, R. Fan 及び G. Ramamurthy が発表した「RRGS: Round Robin Greedy Scheduling for Electric / Optical Terabit Switches」に示されているラウンド・ロビン・グリーディ・スケジューリング (Round Robin Greedy Scheduling: RRGS) アルゴリズムがある。

【0011】

RRGS アルゴリズムを用いたスケジューラでは、分散スケジューリングモジュールがリング状に接続され、隣接した分散スケジューリングモジュール間でメッセージ受渡しを行う。RRGS アルゴリズムでは、各分散スケジューリングモジュールが対象となるタイムスロットの予約 (接続許可決定) を行い、結果の情報を次の分散スケジューリングモジュールに渡す。メッセージ受渡し速度要求条件を緩和するために、RRGS はパイプライン機能を導入している。あるタイムスロットの予約過程は、各分散スケジューリングモジュール間でメッセージ受渡しが一周することにより完了する。また、RRGS アルゴリズムでは、 N 個の分散スケジューリングモジュールが、現在のスロットの少なくとも N スロット先のタイムスロットに関して予約する。更に、RRGS アルゴリズムは、 N 個のタイムスロットに対する予約過程を 1 タイムスロットずつ位相をずらしながら同時に進行させていく。

【0012】

一方、RRGS アルゴリズムの変形として、複数のタイムスロットに対する予

約過程をそれぞれ異なる分散スケジューリングモジュールから同時に開始し、進行させ、同時に終了させるアルゴリズムも考えられる。本アルゴリズムをフレーム化RRGSと呼ぶこととする。

【0013】

図12はRRGS及びフレーム化RRGSを使用する分散型スケジューラによるパケットスイッチの構成を示すブロック図である。同図においては、例としてポート数 $N=4$ の場合を示している。同図において、スケジューラ1は分散型スケジューリングのためのインプットモジュール (Input Module; IM) 10-1~10-4から構成される。各モジュール10- i ($i=1\sim4$) には、フレームの先頭を示すフレームパルス (FP) 21と、スーパーフレームの先頭を示すスーパーフレームパルス (SuperFP) 22が入力される。各モジュール10- i はフレームパルス21もしくはスーパーフレームパルス22に同期して動作する。

【0014】

また、各モジュール10- i にはモジュール識別のための物理番号23が設定される。各入力ポートから接続要求情報11がモジュール10- i に入力され、各モジュール10- i は接続要求の調停の結果、予約（接続許可）を決定し接続許可情報12-1~12-4を出力する。

【0015】

RRGS及びフレーム化RRGSにおいては、隣接分散スケジューリングモジュール間で接続許可情報から入力ポート情報を縮退させた情報（入力ポート情報を参照して作成した情報）である「出力ポート予約済情報」を受け渡すことにより、出力ポートに対する接続要求の競合を回避している。例えば、モジュール10-3は、前段のモジュール10-2から出力ポート予約済情報14-2を出力ポート予約済情報13-3として受信して接続要求の調停に使用する。接続許可情報決定後、出力ポート予約済情報14-3を次段のモジュール10-4に通知する。

【0016】

図13はフレーム化RRGSの一般的な動作例を示した図である。以下、図1

2 及び図 1 3 を参照してフレーム化 R R G S の動作を説明する。例として 4 × 4 の入出力ポートを有するスイッチの動作を示す。

【 0 0 1 7 】

出力ポート予約済情報の転送方向は、モジュールの番号で # 1 → # 2 → # 3 → # 4 → # 1 → # 2 …とする。

【 0 0 1 8 】

各情報の極性を以下のように定義する。各入力ポートの接続要求は、1 が要求あり、0 が要求無しとする。接続許可情報（予約情報）は、1 が許可（予約済）、0 が禁止（未予約）とする。出力ポート予約済情報は、1 が予約済、0 が未予約とする。

【 0 0 1 9 】

また、本例では 4 タイムスロットで 1 フレームを構成しており、フレームパルス 2 1 は 4 タイムスロット周期で入力される。なお、スーパーフレームパルス 2 2 は、本例では使用されない。

【 0 0 2 0 】

図 1 3 において、タイムスロット T S 1 でモジュール 1 0 - 1 ~ 1 0 - 4 にフレームパルス 2 1 が入力されると、モジュール 1 0 - 1 が次フレームのタイムスロット T S 1 に対する入力ポート 1 の接続許可情報を最初に決定する。最初の決定であるので、出力ポート予約済情報は、出力ポート 1 から 4 まで順番に (0 , 0 , 0 , 0) である。入力ポート 1 の接続要求が出力ポート 1 から 4 まで順番に (0 , 1 , 0 , 1) であるとする。モジュール 1 0 - 1 が出力ポート 2 を選択すると、モジュール 1 0 - 1 は次フレームのタイムスロット T S 1 の接続許可情報として出力ポート 2 を記憶し、出力ポート予約済情報を (0 , 1 , 0 , 0) としてモジュール 1 0 - 2 に通知する。

【 0 0 2 1 】

続けてタイムスロット T S 2 でモジュール 1 0 - 2 が次フレームのタイムスロット T S 1 に対する入力ポート 2 の接続許可情報を決定する。モジュール 1 0 - 2 は出力ポート予約済情報 (0 , 1 , 0 , 0) を受信する。入力ポート 2 の接続要求が (0 , 1 , 1 , 1) であるとする、出力ポート 2 に対する割当ては行え

ないため、モジュール10-2は出力ポート2に対する接続要求は拒否し、出力ポート3及び4の接続要求から許可する出力ポートを選択する。ここでポート3を選択すると、モジュール10-2はTS5の接続許可情報として出力ポート3を記憶し、出力ポート予約済情報を(0, 1, 1, 0)としてモジュール10-3に通知する。

【0022】

以下、タイムスロットTS3でモジュール10-3が、タイムスロットTS4でモジュール10-4が次フレームのタイムスロットTS1の接続許可情報を決定する。タイムスロットTS4が終了した時点で各モジュールは次フレームのタイムスロットTS1における接続許可情報を有しているので、次フレームのタイムスロットTS1における4×4の接続許可情報が確定することとなる。

【0023】

更に上記処理手順においては、タイムスロットTS1でモジュール10-1以外のモジュールがそれぞれ異なるタイムスロットに対する「予約」を開始する。例えば、モジュール10-2は次フレームのタイムスロットTS4、モジュール10-3は次フレームのタイムスロットTS3、モジュール10-4は次フレームのタイムスロットTS2に対する「予約」を開始する。

【0024】

各モジュールは当該タイムスロットにて各予約タイムスロットの処理を行い、該当する予約タイムスロットの出力ポート予約済情報をそれぞれ次段のモジュールに転送することにより、各モジュール10-1～10-4とも未動作の時間がないようにスケジューリング処理を実行できる。そして、タイムスロットTS4が終了した時点で各モジュールは次フレームのタイムスロットTS1からTS4までの接続許可情報を有しているので、次フレームのタイムスロットTS1からTS4までの4×4の接続許可情報が確定することとなる。

【0025】

図14は各モジュールの予約（接続許可決定）順序を示す図である。4×4スケジューラの場合を例示している。横軸は時間であり4タイムスロットで1つのフレームを構成している。縦軸にはモジュールの物理番号が並べられている。出

力ポート予約済情報の転送方向は、モジュールの物理番号で # 1 → # 2 → # 3 → # 4 → # 1 → # 2 … である。マトリックス内の数字は予約をする次フレーム内のタイムスロットの番号を指す。

【 0 0 2 6 】

図 1 4 に示されているように、物理番号 # 1 のモジュールは、フレームの先頭（タイムスロット T S 1）で次フレームのタイムスロット T S 1 の予約から開始する。以下同様に、物理番号 # 2 のモジュールはタイムスロット T S 4、物理番号 # 3 のモジュールはタイムスロット T S 3、物理番号 # 4 のモジュールはタイムスロット T S 2 の予約から開始する。

【 0 0 2 7 】

【発明が解決しようとする課題】

上述した従来のアルゴリズムを採用する場合には、以下のような問題点がある。

【 0 0 2 8 】

まず第 1 の問題点は、2 本以上全ポート数未満のポートに均一トラヒックを入力した場合のポート間の予約獲得（接続要求処理）割合が不公平となる点である。RRGS ではパイプライン上で先に予約機会のあるスケジューラの方の予約優先度が高いため、接続機会を獲得する確率が高くなる。従って先に予約機会のあるスケジューラが予約の面で絶対的に有利となる。この特性はモジュールの接続トポロジが固定化されていることに起因しており、この特性のため隣接ポート間の接続機会獲得に関して不公平が生じるという欠点があった。

【 0 0 2 9 】

この点について、上述した図 1 2、図 1 4 を例として説明する。入力ポート 1 と 2 に出力ポート 3 行きのデータが溜まっており、他にはデータが全くないとする。この場合、スケジューラには入力ポート 1 から出力ポート 3 への接続要求と、入力ポート 2 から出力ポート 3 への接続要求が毎タイムスロット送出される。ポート間の公平性を考慮するとこの 2 種類の接続要求は均等に 1 : 1 で処理されるべきである。

【 0 0 3 0 】

しかし図14を参照すると、本スケジューラでは、タイムスロットTS4の予約の時のみ入力ポート2が入力ポート1よりも先に出力ポート3を確保することができ、他のタイムスロットでは入力ポート1が出力ポート3を確保する。従って入力ポート1側が75%、入力ポート2側が25%の処理割合となる。一般にNポート中の隣接2ポートにおいて上記の現象を考えると、パイプラインの上流側と下流側のポートの獲得割合は $N-1:1$ の比率となり、ポート数の増大に伴い不公平の度合いが増大する。本問題は隣接する2ポート以上、 $N-1$ 以下の場合に発生し、隣接ポート数が m の場合最上流のポートとそれ以外のポートの獲得割合は $N-m+1:1:\dots:1$ （項の数は m 個）となる。

【0031】

次に、第2の問題点は、同一フレーム内の処理遅延（接続要求に対する接続許可応答までの時間）に関してポート間に不公平が発生する点である。本例では、各モジュールがフレーム内で予約するタイムスロットの開始順序が固定化されていた。そのため、あるモジュールでは最初に予約機会の与えられるタイムスロットがフレーム内の先頭タイムスロットとなり、他のモジュールでは最初に予約機会の与えられるタイムスロットがフレーム内の最終タイムスロットとなっていた。

【0032】

上述した通り、従来のアルゴリズムでは、パイプライン上で先に予約機会のあるスケジューラの方が予約する確率が高くなるため、先頭に予約機会のあるスケジューラが予約の面で絶対的に有利となる。そのため、各モジュール10の最初に予約機会のあるタイムスロットの送出時刻が固定化されていると、モジュール間で同タイミングに到着した接続要求に対する予約までの時刻差が不公平となり、結果としてセル送出の遅延時間に不公平が発生するという欠点があった。

【0033】

この点について、上述した図12、図14の場合を例として説明する。フレーム先頭であるタイムスロットTS1における各モジュール10-1~10-4の接続予約するタイムスロットに注目すると、モジュール10-1はフレーム先頭で次フレーム・タイムスロットTS1の予約が行われる。モジュール10-2はフ

レーム先頭で次フレーム・タイムスロット T S 4 の予約が行われる。モジュール 1 0 - 1 が最も早いスロットを予約することができ、モジュール 1 0 - 2 は最も遅いタイムスロットを予約することとなり、同タイミングに到着した接続要求に対して常にモジュール 1 0 - 1 の方が、他のモジュール 1 0 - 2 ~ 1 0 - 2 より早いタイムスロットを予約することとなる。

【 0 0 3 4 】

本発明は上述した従来技術の欠点を解決するためになされたものであり、その目的は簡単な付加機能によってポート間の不公平を解消することのできるパケットスイッチ及びパケットスイッチにおけるパケットスイッチング方法を提供することである。

【 0 0 3 5 】

【課題を解決するための手段】

本発明によるパケットスイッチは、ネットワークにおけるデータのフローを制御するスイッチであって、

複数の入力ポートと、

複数の出力ポートと、

前記複数の出力ポートのうちの指定出力ポートへデータを送るように、前記複数の入力ポートのうちの特定の入力ポートを予約し、入力ポートと出力ポートとの接続状態を決定する N 個（N は自然数、以下同じ）の入力ポートスケジューリングモジュールを有するスケジューラと、

を含み、

前記スケジューラは、

タイムスロット単位に、

各入力ポートスケジューリングモジュールが、前段のスケジューリングモジュールから、ある予約タイムスロットの予約状況情報を受信し、

各入力ポートスケジューリングモジュールが、その予約タイムスロットにおける当該入力ポートスケジューリングモジュールからのパケット送出予約可否を決定し、

各入力ポートスケジューリングモジュールが、前段のスケジューリングモジュ

ールから受信した予約状況情報に、自スケジューリングモジュールの予約結果を反映させて、次段のスケジューリングモジュールに送信する手段と、

N個のタイムスロットを単位とするフレームを定義して、前記フレーム時間内で、前記フレームの次のフレーム中のN個のタイムスロットでの予約を行う手段と、

現在のスケジューリングモジュールが、前のスケジューリングモジュールから予約状況情報を受信する手段と、

前記現在のスケジューリングモジュールが、前記複数の出力ポートのうちの1つにアクセスしようとする将来のタイムスロットを予め次のフレーム内の特定のタイムスロットに決定する手段と、

前記複数の出力ポートのうちの1つを前記将来のタイムスロットでの送信用に予約するように選択する手段と、

前記将来のタイムスロットが他のスケジューリングモジュールによって予約済みかどうかを判断する手段と、

前記将来のタイムスロットが他のスケジューリングモジュールによって予約されていない場合には、前記将来のタイムスロットを予約し、前記将来のタイムスロットが予約されたことを示す情報を前記予約状況情報に入れる手段と、

前記予約状況情報を次のスケジューリングモジュールに渡す手段と、

タイムスロットでの予約過程の観点から見た場合、

前記予約過程が、フレームの先頭で同時に開始され、フレーム内で同時にパイプライン処理により進行し、

フレームの末端で同時に完了し、

前記入力ポートスケジューリングモジュールが、フレームの先頭で同時に開始する複数の前記予約過程それぞれにおいて、次のフレーム内の各々異なる予約タイムスロットを対象として処理を開始する手段と、

前記複数のスケジューリングモジュールによる前記予約の順序を変更する予約順序変更手段とを含み、

この変更された順序で前記複数のスケジューリングモジュールそれぞれが各フレーム毎に次のフレームの packets について出力すべきポートを予約するように

したことを特徴とする。

【0036】

また、前記複数のスケジューリングモジュールは他のモジュールとの論理的接続順序に対応した順序で前記予約を実行するものであり、前記予約順序変更手段は前記複数のスケジューリングモジュールの接続トポロジを変更する。前記予約順序変更手段は、前記複数のスケジューリングモジュールの論理的接続状態を変更するようにスイッチング動作するスイッチと、このスイッチのスイッチング動作を制御する制御データを記憶したテーブルとを含む。前記複数のスケジューリングモジュールと前記スイッチ間の物理的接続は、電氣的接続か光学的接続とする。

【0037】

さらに、前記テーブルは、前記複数のスケジューリングモジュールそれぞれの内部に設けられているか、前記複数のスケジューリングモジュールに対して共通に設けられているものとする。そして、前記制御データは、前記複数のスケジューリングモジュール各々がフレーム内の先頭において予約を開始するタイムスロットを各フレーム毎に変更できるように前記スイッチのスイッチング動作を制御するデータであるか、前記予約を開始する予約開始スロットの前記複数のスケジューリングモジュールによる使用頻度を均等とするスケジューリングを実現するデータであるか、前記予約を開始する予約開始スロットの前記複数のスケジューリングモジュールによる使用順序及び使用頻度を変更可能とするスケジューリングを実現するデータであるものとする。

【0038】

本発明によるパケットスイッチング方法は、N個の入力ポートスケジューリングモジュールを有するスイッチのスケジューラにて使用する、複数の出力ポートのうちの指定出力ポートヘデータを送るように、複数の入力ポートのうちの特定の入力ポートを予約し、入力ポートと出力ポートとの接続状態を決定するパケットスイッチング方法であって、

各入力スケジューリングモジュールが、前段のスケジューリングモジュールからあるタイムスロットの予約状況情報を受信するステップと、

各スケジューリングモジュールが、その予約タイムスロットにおける当該入力ポートスケジューリングモジュールからのパケット送出予約可否を決定するステップと、

各入力スケジューリングモジュールが、前段のスケジューリングモジュールから受信した予約状況情報に、自スケジューリングモジュールの予約結果を反映させて、次段のスケジューリングモジュールに送信するステップと、

を含み、

前記スケジューラは、

現在のスケジューリングモジュールが、前のスケジューリングモジュールから予約状況情報を受信するステップと、

前記現在のスケジューリングモジュールが、前記複数の出力ポートのうちの1つにアクセスしようとする将来のタイムスロットを予め次のフレーム内のタイムスロットに決定するステップと、

前記複数の出力ポートのうちの1つを前記将来のタイムスロットでの送信用に予約するように選択するステップと、

前記将来のタイムスロットが他のスケジューリングモジュールによって予約済みかどうかを判断するステップと、

前記将来のタイムスロットが他のスケジューリングモジュールによって予約されていない場合には、前記将来のタイムスロットを予約し、前記将来のタイムスロットが予約されたことを示す情報を前記予約状況情報に入れるステップと、

前記予約状況情報を次のスケジューリングモジュールに渡すステップと、

を実行し、

タイムスロットでの予約過程の観点から見た場合、

前記予約過程が、フレームの先頭で同時に開始され、フレーム内で同時にパイプライン処理により進行し、

フレームの末端で同時に完了し、

前記入力ポートスケジューリングモジュールが、フレームの先頭で同時に開始する複数の前記予約過程それぞれにおいて、次のフレーム内の各々異なる予約タイムスロットを対象として処理を開始し、

さらに、前記複数のスケジューリングモジュールによる前記予約の順序を変更し、この変更された順序で前記複数のスケジューリングモジュールそれぞれが各フレーム毎に次のフレームの packets について出力すべきポートを予約するようにしたことを特徴とする。前記複数のスケジューリングモジュールは他のモジュールとの論理的接続順序に対応した順序で前記予約を実行するものであり、前記予約順序の変更においては前記複数のスケジューリングモジュールの接続トポロジを変更することを特徴とする。

【0039】

また、前記予約の順序の変更においては、前記複数のスケジューリングモジュールの論理的接続状態を変更するようにスイッチング動作するスイッチと、このスイッチのスイッチング動作を制御する制御データを記憶したテーブルとを用いて予約順序を変更することを特徴とする。前記複数のスケジューリングモジュールと前記スイッチ間の物理的接続は、電気的接続か、光学的接続とする。前記テーブルは、前記複数のスケジューリングモジュールそれぞれの内部に設けられているか、前記複数のスケジューリングモジュールに対して共通に設けられているものとする。

【0040】

前記制御データは、前記複数のスケジューリングモジュール各々がフレーム内の先頭において予約を開始するタイムスロットを各フレーム毎に変更できるように前記スイッチのスイッチング動作を制御するデータであるか、前記予約を開始する予約開始スロットの前記複数のスケジューリングモジュールによる使用頻度を均等とするスケジューリングを実現するデータであるか、前記予約を開始する予約開始スロットの前記複数のスケジューリングモジュールによる使用順序及び使用頻度を変更可能とするスケジューリングを実現するデータであるものとする。

【0041】

要するに本発明では、入力バッファの構成として仮想出力キューを有し、データスイッチング素子としてクロスバースイッチを有する、入力バッファ型の packets スイッチを実現しているのである。そして、各モジュールは、入力バッファ

の各VOQからの接続要求に対してあるタイムスロットのクロスバスイッチの入力ポートと出力ポートの接続情報を決定する。

【0042】

前述の通り、RRGS、フレーム化RRGSには2つの不公平性が存在する。すなわち、第1に、RRGS、フレーム化RRGS共通の問題で、隣接ポート間における接続予約機会に関する不公平性である（第1の不公平性）。また、第2に、フレーム化RRGSについての問題で、ポート間における接続要求に対する接続許可応答までの同一フレーム内の処理遅延に関する不公平性である（第2の不公平性）。

【0043】

そこで、本発明では、第1の不公平性の解消のために、モジュール間の接続トポロジが変更できる構成をとる。装置としては、モジュールの間で受け渡しをしている出力ポート予約済情報をモジュール外の切替スイッチに入力し、出力先を変更できるようにする。モジュールと切替スイッチは、モジュールが複数タイムスロットの接続許可処理を行っているフレームを単位に同期して動作し、フレーム単位に接続トポロジを変更し全接続トポロジのパターンが出現するように動作する。

【0044】

接続トポロジが変更されることにより、様々な隣接ポートの組み合わせが出現し、固定化されていた入力ポートに関する優先度がシャッフルされることになり、入力ポートの予約機会に関する不公平性が解消される。また、特定の接続トポロジのパターンのみを用意して使用したり、パターンの適用頻度に重み付けを与えて各パターンの適用割合を変化させることにより、特定の入力ポートもしくは複数の特定の入力ポートに対する優先ポート割当ての頻度を上げる制御を実現することができる。

【0045】

そして、第2の不公平性の解消のためには、モジュール内の処理フレームでの予約タイムスロットの処理順序を、フレーム単位で変更する。装置としては、各モジュールがフレームに同期して、各モジュール毎に予約タイムスロットの順序

パターンを変更する。予約タイムスロットの順序パターンを各モジュールが同期して変更することにより、固定化していた接続要求に対する接続許可応答までの遅延時間の平均値が変更できるようになる。

【0046】

これにより、平均値を平準化することが実現でき、入力ポートの処理遅延に関する不公平性が解消される。また、特定の予約タイムスロットの順序パターンのみを用意して使用したり、パターンの適用頻度に重み付けを与えて各パターンの適用割合を変化させることにより、特定の入力ポートもしくは特定の複数の入力ポートに対する処理遅延に関する優先制御を実現することができなる。

【0047】

【発明の実施の形態】

次に、本発明の実施の一形態について図面を参照して説明する。なお、以下の説明において参照する各図においては、他の図と同等部分には同一符号が付されている。

【0048】

図2を参照すると、本発明の第1の実施の形態としての各モジュールの接続トポロジのタイプが示されている。本形態では第1の問題点を解決する方法を示している。例ではポート数 $N=4$ の場合を示している。ポート数が N の場合接続トポロジの数は N の円順列となり、 $(N-1)$ の階乗 $((N-1)!)$ となる。スケジューラはこの $(N-1)!$ 分の接続トポロジパターンを用意し、各接続トポロジタイプを均等に発生させてモジュールの接続を変更することにより、隣接モジュール間のポート予約に関する不公平性を解消することができる。本例の場合、 $(4-1)! = 6$ 通りの接続トポロジタイプが存在する。この6通りの接続トポロジタイプを均等に発生させてモジュールの接続を変更する。

【0049】

論理的モジュール番号を a から d までとし、実モジュール番号を1から4までとする。論理的モジュール番号で $a \rightarrow b \rightarrow c \rightarrow d \rightarrow a$ の順に出力ポート予約済情報の転送をする。6通りの接続トポロジタイプにおける実モジュール番号の割当ては図2に示されているようになる。例えば、Type #1では物理番号1のモ

ジュールの論理番号は a となり、物理番号 2 の論理番号は b、物理番号 3 の論理番号は c、物理番号 4 の論理番号は d となる。

【 0 0 5 0 】

図 3 を参照すると、本発明の第 1 の実施の形態としての各モジュールの予約（接続許可決定）順序が示されている。縦軸、横軸、マトリックス内の数字の内容は図 1 4 の場合と同様である。例ではポート数 $N = 4$ の場合を示している。複数タイムスロットからフレームを構成しており、1 フレームはポート数から 4 タイムスロットとなる。また複数フレームからスーパーフレームを構成している。フレーム 1 からフレーム 6 までの 6 フレームで前述の接続トポロジタイプ 1 から 6 までを適用している。この 6 フレームをまとめて 1 つのスーパーフレームが構成される。6 通りの接続トポロジタイプを順番にフレーム単位に適用することにより、隣接モジュール間のポート予約に関する不公平性を解消している。なお、各フレームとも論理番号 a の予約開始タイムスロットが 1、b が 4、c が 3、d が 2 となる接続許可情報予約開始パターンにて例を示している（これを Type # A の接続許可情報予約開始パターンとする）。

【 0 0 5 1 】

図 1 を参照すると、本発明の第 1 の実施例によるスケジューラの構成が示されている。本例においてもポート数 $N = 4$ の場合を示す。スケジューラ 1 はポート数分のモジュール 1 0 - 1 ~ 1 0 - 4 と、切替スイッチ 3 0 とから構成されている。

【 0 0 5 2 】

各モジュール 1 0 - i には、フレームの先頭を示すフレームパルス (FP) 2 1 と、スーパーフレームの先頭を示すスーパーフレームパルス (SuperFP) 2 2 が入力される。また、各モジュール 1 0 - i にはモジュール識別のための物理番号 2 3 が設定される。さらに、各モジュール 1 0 - i には、接続要求情報 1 1 と出力ポート予約済情報 1 3 が入力される。

【 0 0 5 3 】

モジュール 1 0 - i は、接続要求の調停を行って接続許可（予約）を決定し、接続許可情報 1 2 と更新された出力ポート予約済情報 1 4 を出力する機能を有し

ている。各モジュール 1 0 は出力ポート予約済情報 1 4、及び切替情報 2 0 を切替スイッチ 3 0 に入力し、切替スイッチ 3 0 は切替情報 2 0 を基に出力ポート予約済情報 1 3 をスイッチングして各モジュール 1 0 に出力する。

【0 0 5 4】

ここで、図 1 中の各モジュール 1 0 - i の構成例について図 4 を参照して説明する。同図を参照すると、モジュール 1 0 - i は、アロケータ 1 5 と、接続許可記憶部 1 6 と、接続許可記憶制御部 1 7 と、パタン記憶部 1 8 と、パタン読出制御部 1 9 とを含んで構成されている。

【0 0 5 5】

アロケータ 1 5 は、接続要求情報 1 1 と出力ポート予約済情報 1 3 とから、本モジュールで管理する入力ポートに対する出力ポートの接続許可情報 1 2 を決定する。決定のためのアルゴリズムは公知のアルゴリズムを適用する。

【0 0 5 6】

接続許可記憶部 1 6 は、アロケータ 1 5 にて決定した接続許可情報 1 2 を使用するタイムスロットの時刻まで記憶しておく機能を有する。この接続許可記憶部 1 6 は、図 5 に示されているように、接続許可情報を記憶するためのメモリ 1 6 0 を含んで構成されている。

【0 0 5 7】

接続許可記憶制御部 1 7 は、フレームパルス 2 1 に同期して、パタン記憶部 1 8 からの接続許可情報の予約順序パタンと、モジュール識別のための物理番号 2 3 から、当該モジュールにおける接続許可情報の予約順序パタンとを決定し、タイムスロット毎の接続許可情報 1 2 の書込読出順序を制御する。この接続許可記憶制御部 1 7 は、図 5 に示されているように、接続許可記憶部 1 6 内のメモリ 1 6 0 に対する書込アドレスを生成するための書込アドレスカウンタ 1 7 0 と、同じく読出アドレスを生成するための読出アドレスカウンタ 1 7 1 と、ロードデータ生成部 1 7 2 とを含んで構成されている。

【0 0 5 8】

ロードデータ生成部 1 7 2 は、接続許可情報予約開始パタンと接続トポロジタイプと物理番号 2 3 から接続許可情報予約開始値を決定する。書込アドレスカウ

ンタ 170 は、接続許可情報予約開始値をロードデータ (Load Data) とし、フレームパルスを入力 (Load) とする。また、読出アドレスカウンタ 171 は、フレームパルスを入力 (Load) とする。これらのカウンタ 170 及び 171 は、共に、タイムスロット時間を 1 周期とする図示せぬクロックに応じてカウント動作を行う。そして、そのカウント値を書込アドレス及び読出アドレスとして接続許可記憶部 16 内のメモリ 160 に入力し、接続許可情報の書込み及び読出しを行う。

【0059】

パタン記憶部 18 は、接続許可情報 12 の出力順序パタンを決定するためのパタン情報を記憶している。このパタン記憶部 18 は、図 6 (A) に示されているように、パタン番号を入力とし、接続トポロジタイプ及び接続許可情報予約開始パタンを出力とするパタンテーブル 180 を含んで構成されている。同図 (B) を参照すると、本発明の第 1 の実施例によるパタンテーブル 180 の内容が示されている。パタンテーブル 180 には、0 から 5 までの各パタン番号に対応して Type # 1 から # 6 の接続トポロジタイプが保持されている。また、本例では各パタン番号に共通して Type # A の接続許可情報予約開始パタンが保持されている。このような内容のパタンテーブル 180 を設けることにより、パタン記憶部 18 は、図 3 に示されている、接続トポロジパタン毎の接続許可情報の予約順序パタンを出力することができる。

【0060】

パタン読出制御部 19 は、スーパーフレームパルス 22 に同期して、フレーム毎に接続許可情報 12 の出力順序のパタンの読出を決定し、接続許可記憶制御部 17 にパタンを通知する。また、パタン読出制御部 19 は、切替スイッチ 30 に対して接続トポロジの切替情報 20 を通知する。このパタン読出制御部 19 は、図 7 に示されているように、フレームを周期とするクロック (CLK) に同期して動作するカウンタ 190 と、接続トポロジタイプを切替情報 20 に変換する切替情報変換部 191 とを含んで構成されている。

【0061】

カウンタ 190 のカウント値は、クロックの入力に応じて増加していき、パタ

ン番号としてパタン記憶部 18 に送出される。そして、このカウント値がスーパーフレームパルス 22 の遷移タイミングでリセットされることにより、同じパタン番号が繰返しパタン記憶部 18 に送出されることになる。また、パタン記憶部 18 から送出される接続トポロジタイプが切替情報変換部 191 に入力されると、切替情報変換部 191 はそのタイプに対応する切替情報 20 を出力する。

【0062】

図 1 中の切替スイッチ 30 の内部構成例について図 8 を参照して説明する。同図に示されているように、切替スイッチ 30 は、4 つの入力のうちの 1 つを出力する 4 : 1 セレクタ (SEL) 101 ~ 104 と、セレクタ 101 ~ 104 に対応して設けられ対応するセレクタに入力するデータを一時保持するフリップフロップ (以下、F/F と略す) 100a ~ 100d と、セレクタ 101 ~ 104 に対応して設けられ対応するセレクタから出力されるデータを一時保持する F/F 101b, 102b, 103b, 104b と、セレクタ 101 ~ 104 の制御端子 CTL に与える切替情報を一時保持する F/F 101a, 102a, 103a, 104a とを含んで構成されている。

【0063】

また、同図中の F/F 100a ~ 100d に入力するデータが印加されるポート IS1UP ~ ポート IS4UP には、上述した出力ポート予約済み情報 14-1 ~ 14-4 が入力される。ポート IUP1SW ~ ポート IUP4SW には、上述した切替情報 20-1 ~ 20-4 が入力される。切替情報 20-1 ~ 20-4 は、4 : 1 セレクタ 101 ~ 104 が各々異なる入力を選択するような情報として、ポート IS1UP ~ ポート IS4UP に入力される。ポート OS1UP ~ ポート OS4UP からは、F/F 101b, 102b, 103b, 104b に保持されている出力ポート予約済み情報 13-1 ~ 13-4 が出力される。

【0064】

なお、同図中の各 F/F は、ポート ISYSCCLK に印加されているシステムクロックに同期して動作するものとする。また、同図中のポートIRSTB に印加されるリセット信号に応答して各 F/F はクリア状態になるものとする。

【0065】

かかる構成からなる切替スイッチ 30 においては、各モジュール 10-1~10-4 から出力される 4 種類の出力ポート予約済み情報が 4:1 セレクタ 101~104 にそれぞれ入力される。そして、これらセレクタ 101~104 それぞれにおいて 1 種類の出力ポート予約済み情報が選択されて出力される。この選択された出力ポート予約済み情報は、各モジュール 10-1~10-4 に入力される。

【0066】

次に、本実施例のスケジューラの動作について、各モジュール 10-i が図 3 の順序に従ってポートの予約処理を行う場合を説明する。

【0067】

フレームパルス 21 は 4 タイムスロット周期、スーパーフレームパルス 22 は 6 フレーム周期 (24 タイムスロット周期) で入力される。図 3 のフレーム 1 のタイムスロット 1 において、図 4 に示されているモジュール 10-i にフレームパルス 21、スーパーフレームパルス 22 が入力されると、パタン読出制御部 19 は、図 2 の接続トポロジタイプ 1 をパタン記憶部 18 から接続許可記憶制御部 17 にセットする。また、パタン読出制御部 19 は接続トポロジタイプ 1 の切替情報を出力する。各モジュールは自モジュールを識別するための物理番号 23 と設定された接続トポロジタイプから現フレームでの接続許可情報の予約順序パタンを選択する。例えば物理番号 23 に「1」が設定されているモジュール 10-1 は図 3 の接続許可情報の予約順序にて物理番号 1、フレーム 1 のパタンを選択する。

【0068】

図 1 中の切替スイッチ 30 はモジュール 10 からの切替情報 20 に基づいて、各モジュール 10-1~10-4 からの出力ポート予約済み情報の接続を、図 2 の接続トポロジタイプ 1 となるように接続する。図 1 中の各モジュール 10-1~10-4 は、図 3 に示されているフレーム 1 の区間で接続トポロジタイプ 1 の場合の接続許可情報予約順序に従って予約を実施する。

【0069】

図 3 のフレーム 1 終了時にはフレーム 2 の 4 個のタイムスロットにおける接続

許可情報 12 が確定し、図 4 中の接続情報記憶部 16 に記憶されている。図 1 中の各モジュール 10-1 ~ 10-4 は、図 3 のフレーム 2 のタイムスロット 1 から 4 の区間で、決定済の接続許可情報を使用するタイムスロットに応じて図 4 中の接続情報記憶部 16 から接続許可情報 12 を読出して、クロスバースイッチと入力ポートに通知し、データのスイッチングを実行する。

【0070】

上記の接続許可情報の通知と同時に次フレームの接続許可情報の決定を行う。図 3 のフレーム 2 のタイムスロット 1 においてフレームパルス 21 が入力されると、図 4 中のパタン読出制御部 19 は、図 2 の接続トポロジタイプ 2 をパタン記憶部 18 から接続許可記憶制御部 17 にセットする。また、パタン読出制御部 19 から接続トポロジタイプ 2 の切替情報を出力する。各モジュール 10-i は自モジュールを識別するための物理番号 23 と設定された接続トポロジタイプ 2 から現フレームでの接続許可情報の予約順序パタンを選択する。

【0071】

図 1 中の切替スイッチ 30 はモジュール 10-1 ~ 10-4 からの切替情報 20 に基づいて、各モジュール 10-1 ~ 10-4 からの出力ポート予約済情報の接続を図 2 の接続トポロジタイプ 2 となるように接続する。図 1 中の各モジュール 10-1 ~ 10-4 は、図 3 のフレーム 2 の区間で接続トポロジタイプ 2 の場合の接続許可情報予約順序に従って予約を実施する。以下同様に、接続トポロジの変更と、次フレームの予約と、接続許可情報の通知を行う。タイプ 6 の接続トポロジによる予約が実行された後には、フレームパルス 21、スーパーフレームパルス 22 が入力される。モジュール 10 は接続トポロジタイプをタイプ 1 に戻し、予約処理を継続する。

【0072】

上述したように、切替スイッチ 30 によって各モジュール 10-1 ~ 10-4 間の接続トポロジを変更し、全接続トポロジのパタンを均一に出現させることにより、固定化されていた入力ポートの予約機会に関する優先度がシャッフルされることになり、入力ポートの予約機会に関する不公平性が解消される。

【0073】

上述した第1の実施例においては、切替スイッチ30における出力ポート予約済情報のスイッチのための切替情報を各モジュール10-1~10-Nから通知する形態について説明した。これにより、各モジュール10-iを同一構成とし、各モジュールにフレームパルス21及びスーパーフレームパルス22を入力し、更にモジュール識別のための物理番号23を外部から入力するだけで、容易にパケットスイッチを構成できる。このように、各モジュール10-iを同一構成とすることができるので、モジュールの量産に適し、製造コストの低減が図れる。なお、この切替情報のある特定のモジュール10-k (kは自然数、かつ、 $1 \leq k \leq N$) から切替スイッチ30に通知する形態も考えられる。

【0074】

また、モジュール10-iにて、接続トポロジパターンを読み出す方法としては、各パターンをラウンドロビンにて読み出すことにより均一にパターンを使用する方法がある。その他、適用したい順番に各パターンをパターン記憶部18に記憶させておくという方法も考えられる。

【0075】

さらにまた上述した実施例では、モジュール10-iにフレームパルス21及びスーパーフレームパルス22を入力し、スーパーフレームパルス22に同期して各モジュール10-iの接続トポロジタイプの変更順序の同期をとり、フレームパルス21に同期して接続トポロジタイプ及び接続許可情報の予約順序パターンを更新している。この構成以外に接続トポロジタイプや接続許可情報の予約順序パターンをモジュール10-iの外部から入力する構成も考えられる。つまり、図9に示されているように、接続トポロジタイプや接続許可情報の予約順序パターンを発生する情報発生部140を設ければ良い。ただし、同期をとるために、図示せぬ信号線によってフレームパルス21及びスーパーフレームパルス22を情報発生部140に入力する必要がある。

【0076】

なお、上述した第1の実施例では全接続トポロジのパターンを均一に発生させたが、逆に特定の接続トポロジパターンの出現頻度を高めることもできる。この場合、特定の入力ポートに対する予約機会に関する優先度を高めることができるとい

う効果がある。

【0077】

次に、本発明の第2の実施例について説明する。図10には、本実施例における、各モジュール10-iの予約（接続許可決定）順序が示されている。本形態では第2の問題点を解決する方法を示している。縦軸、横軸、マトリックス内の数字の内容は図14の場合と同様である。例では同じくポート数 $N=4$ の場合を示している。複数タイムスロットからフレームを構成しており、1フレームはポート数から4タイムスロットとなる。また複数フレームからスーパーフレームを構成している。ポート数が N の場合、モジュール10-iがフレーム内の先頭において予約を開始するタイムスロットの場合の数（位相タイプと呼ぶ）は N 通りとなる。なお、各フレームとも、接続トポロジパターンとしてはType #1のパターンにて例を示している。すなわち、物理番号1のモジュールの論理的モジュール番号がa、物理番号2のモジュールの論理的モジュール番号がb、物理番号3のモジュールの論理的モジュール番号がc、物理番号3のモジュールの論理的モジュール番号がdとなる例を示している。

【0078】

本例の場合、位相タイプとしてType #AからType #Dまでの4通りの場合が存在する。例えば、Type #Aでは、論理的モジュール番号aの予約開始タイムスロットが1となり、論理番号bの予約開始タイムスロットが4、論理番号cの予約開始タイムスロットが3、論理番号dの予約開始タイムスロットが2となる。従って、Type #Aでは、物理番号1のモジュールの予約開始タイムスロットが1となり、物理番号2の予約タイムスロットが4、物理番号3の予約タイムスロットが3、物理番号4の予約タイムスロットが2となる。フレーム1から4までで前述の位相タイプAからDまでを適用している。この4フレームをまとめて1つのスーパーフレームが構成される。4通りの位相タイプを順番にフレーム単位に適用することにより、各モジュール間の遅延に関する不公平性を解消している。

【0079】

本実施例におけるスケジューラの全体構成は、図12に示されている構成と同

様である。ただし、図12中の各モジュール10-1~10-4は、図4に示されている構成であるものとする。

【0080】

各モジュール10-iには、フレームの先頭を示すフレームパルス(FP)と、スーパーフレームの先頭を示すスーパーフレームパルス(SuperFP)が入力される。モジュール10-iの切替情報20は本実施例では使用しない。

【0081】

モジュール10-i内のパタン記憶部18は、接続許可情報の予約順序のパタンを記憶しておく。図6(C)を参照すると、本発明の第2の実施例によるパタンテーブルの内容が示されている。パタンテーブル180には、0から3までの各パタン番号に対応してType #Aから#Dの接続許可情報予約開始パタンが保持されている。また、本例では各パタン番号に共通してType #1の接続トポロジタイプが保持されている。このような内容のパタンテーブル180を設けることにより、パタン記憶部18は、図6(C)に示されているパタンを記憶している。このような内容のパタンテーブル180を設けることにより、パタン記憶部18は、図10に示されている位相パタン毎の接続許可情報の予約順序パタンを出力することができる。

【0082】

次に、本発明の第2の実施例のスケジューラの動作について説明する。図12の各分散スケジューリングモジュール10-1~10-Nが図10に示されている順序に従ってポートの予約処理を行う場合を説明する。なお、フレームパルス21は4タイムスロット周期、スーパーフレームパルス22は4フレーム周期(16タイムスロット周期)で入力されるものとする。

【0083】

図10中のフレーム1のタイムスロット1において、図4に示されているモジュール10-iにフレームパルス21、スーパーフレームパルス22が入力されると、パタン読出制御部19は、図10中の位相タイプAをパタン記憶部18から接続許可記憶制御部17にセットする。各モジュールは自モジュールを識別するための物理番号23と設定された位相タイプから現フレームでの接続許可情報

の予約順序パターンを選択する。例えば物理番号23に1が設定されているモジュール10-1は図10中の接続許可情報の予約順序にて物理番号1、フレーム1のパタンを選択する。

【0084】

各モジュール10-iは、フレーム1の区間で位相タイプAの場合のスケジューリング順序に従ってフレーム化RRGSによるスケジューリングを実施する。図10中のフレーム1終了時にはフレーム2の4個のタイムスロットにおける接続許可情報が確定し、図4中の接続情報記憶部16に記憶されている。各モジュールは、フレーム2のタイムスロット1から4にかけて決定済の接続許可情報をタイムスロットに応じて図4中の接続情報記憶部16から接続許可情報12を読み出して、クロスバースイッチと入力ポートに通知し、データ転送を実行する。

【0085】

上記接続許可情報の通知と同時に次フレームの接続許可情報の決定を行う。図10中のフレーム2のタイムスロット1においてフレームパルス21が入力されると、図4中のパタン読出制御部19は、図10中の位相タイプBをパタン記憶部18から接続許可記憶制御部17にセットする。各モジュールは自モジュールを識別するための物理番号23と設定された位相タイプBから現フレームでの接続許可情報の予約順序パターンを選択する。

【0086】

各モジュール10-1~10-4はフレーム2の区間で位相タイプBの場合の接続許可情報予約順序に従って予約を開始する。

【0087】

以下同様に、位相タイプの変更と、次フレームの予約と、接続許可情報の通知を行う。タイプDの位相タイプによる予約が実行された後には、フレームパルス21、スーパーフレームパルス22が入力される。モジュール10は位相タイプをタイプAに戻し、予約処理を継続する。

【0088】

以上のように、各モジュール10-1~10-N間の予約開始タイムスロットの位相パターンを変更し、全位相パターンが均一に出現することにより、固定化され

ていた入力ポートの平均遅延がシャッフルされることになる。よって、入力ポートの平均遅延に関する不公平性が解消される。

【 0 0 8 9 】

入力ポートの平均遅延に関する不公平性の解消には、各モジュール 1 0 - 1 ~ 1 0 - N の接続トポロジを変更する必要がないので、第 1 の実施例のようなモジュール 1 0 - i 外で出力ポート予約済情報をスイッチする機能は必要ない。従って外部に回路を追加しなくても、第 1 の実施例と同じモジュール 1 0 - i の構成で実現できる。

【 0 0 9 0 】

第 1 の実施例の場合と同様に、位相パタンの読出に関して各パタンをラウンドロビンにて読出すことにより均一にパタンを使用する方法がある。その他、適用したい順番に各パタンをパタン記憶部 1 8 に記憶させておくという方法も考えられる。

【 0 0 9 1 】

また、上述の実施例では各モジュール 1 0 - i にフレームパルス 2 1 及びスーパーフレームパルス 2 2 を入力し、スーパーフレームパルス 2 2 に同期して各モジュール 1 0 - i の位相タイプの変更順序の同期をとり、フレームパルス 2 1 に同期して位相タイプ及び接続許可情報の予約順序パタンを更新していた。第 1 の実施例と同様に、本構成以外に位相タイプや接続許可情報の予約順序パタンをモジュール 1 0 - i の外部から入力する方法も考えられる。

【 0 0 9 2 】

更に、上記実施の形態では全位相パタンを均一に発生させたが、逆に特定の位相パタンの出現頻度を高めて、特定の入力ポートの平均遅延を小さくすることもできる。

【 0 0 9 3 】

なお、第 1 の実施例の構成では、平均遅延の公平化と入力ポートの予約機会に関する公平化を同時に実現することができる。

【 0 0 9 4 】

モジュール 1 0 - i の接続トポロジは、各モジュール 1 0 - 1 ~ 1 0 - N 間の

予約開始タイムスロットの位相を固定にした場合に相当する。図 3 及び図 1 0 を例にとって説明すると、ポート数 $N=4$ の場合、図 3 に示されている 6 通りの各モジュールの予約順序は、図 1 0 に示されている位相タイプ A で接続変更を実施した場合に対応する。従って図 1 0 のタイプ B, C, D の場合に接続変更を実施したパターンを考慮すると、全部で 24 通りのパターンが存在することとなる。これはスケジューラ 1 におけるモジュールの順列に他ならない。すなわち、モジュールの接続トポロジタイプと予約開始タイムスロットの位相タイプとの直積（掛け算）がモジュールの順列となる。

【0095】

第 1 の実施例の構成で、接続トポロジパターンを記憶しているパターン記憶部 1 8 に接続トポロジパターン（6 通り）に代わってモジュールの順列（24 通り）を記憶し、スーパーフレームパルス 2 2 を 24 フレーム（96 タイムスロット）の整数倍周期で入力して、パターン読出制御部 1 9 が 24 通りの読出を実行するように変更することによって、入力ポートの予約機会に関する公平化と平均遅延の公平化とを同時に解消できる。図 6（D）を参照すると、上述の本発明の他の実施例によるパターンテーブル 1 8 0 の内容が示されている。パターンテーブル 1 8 0 には、0 から 23 までの各パターン番号に対応して、Type # 1 から # 6 の接続トポロジタイプと Type # A から # D の接続許可情報予約開始パターンの全組み合わせが保持されている。

【0096】

【発明の効果】

以上説明したように本発明は、接続トポロジを変更することにより、様々な隣接ポートの組み合わせが出現し、固定化されていた入力ポートに関する優先度がシャッフルされることになり、入力ポートの予約機会に関する不公平性が解消できるという効果がある。

【0097】

また、各モジュール内の処理フレームでの予約タイムスロットの処理順序を、フレーム単位で変更することにより、固定化していた接続要求に対する接続許可応答までの遅延時間の平均値が変更できるようになり、平均値を平準化すること

が実現できて、入力ポートの処理遅延に関する不公平性が解消できるという効果がある。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施例によるパケットスイッチの構成を示すブロック図である。

【図 2】

本発明の第 1 の実施例によるパケットスイッチの各モジュールの接続トポロジのタイプを示す図である。

【図 3】

第 1 の実施例の各モジュールの予約（接続許可決定）順序を示す図である。

【図 4】

図 1 中の各モジュールの構成例を示すブロック図である。

【図 5】

図 4 中の接続許可記憶部及び接続許可記憶制御部の内部構成例を示す図である。

【図 6】

図（A）は図 4 中のパタン記憶部の構成例を示す図、図（B）は本発明の第 1 の実施例による図（A）中のパタンテーブルの内容例を示す図、図（C）は本発明の第 2 の実施例による図（A）中のパタンテーブルの内容例を示す図、図（D）は本発明の他の実施例による図（A）中のパタンテーブルの内容例を示す図である。

【図 7】

図 4 中のパタン読出制御部の内部構成例を示す図である。

【図 8】

図 1 中の切替スイッチの内部構成例を示す図である。

【図 9】

本発明の第 2 の実施例によるパケットスイッチの構成を示すブロック図である。

【図 1 0】

本発明の第 2 の実施例によるパケットスイッチの各モジュールの予約（接続許可決定）順序を示す図である。

【図 1 1】

V O Q 使用の入力バッファ型パケットスイッチの一般的な構成を示す図である。

【図 1 2】

一般的なパケットスイッチの構成を示す図である。

【図 1 3】

フレーム化 R R G S の接続許可情報の決定方法を示す図である。

【図 1 4】

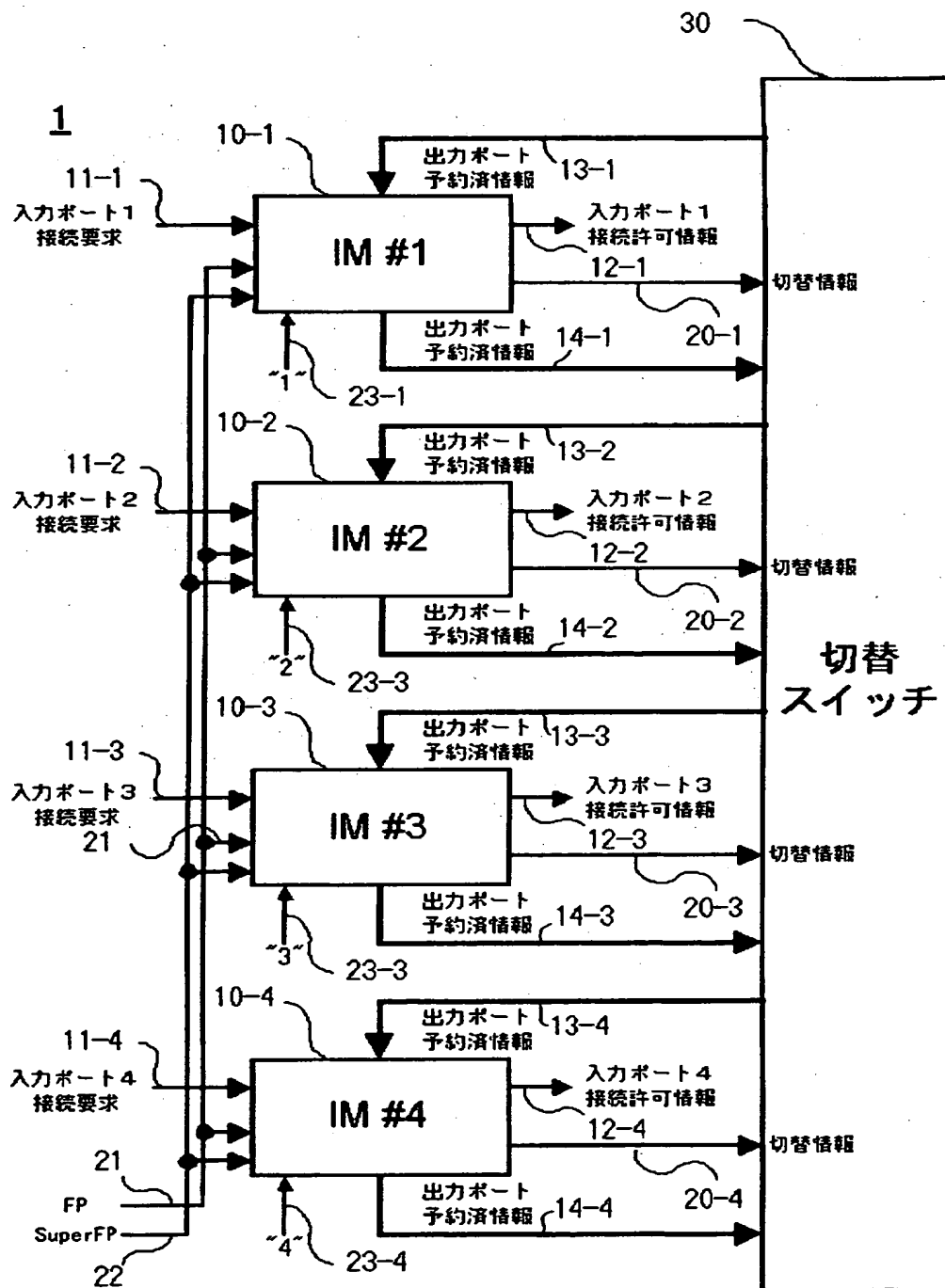
従来のフレーム化 R R G S における各モジュールの予約（接続許可決定）順序を示す図である。

【符号の説明】

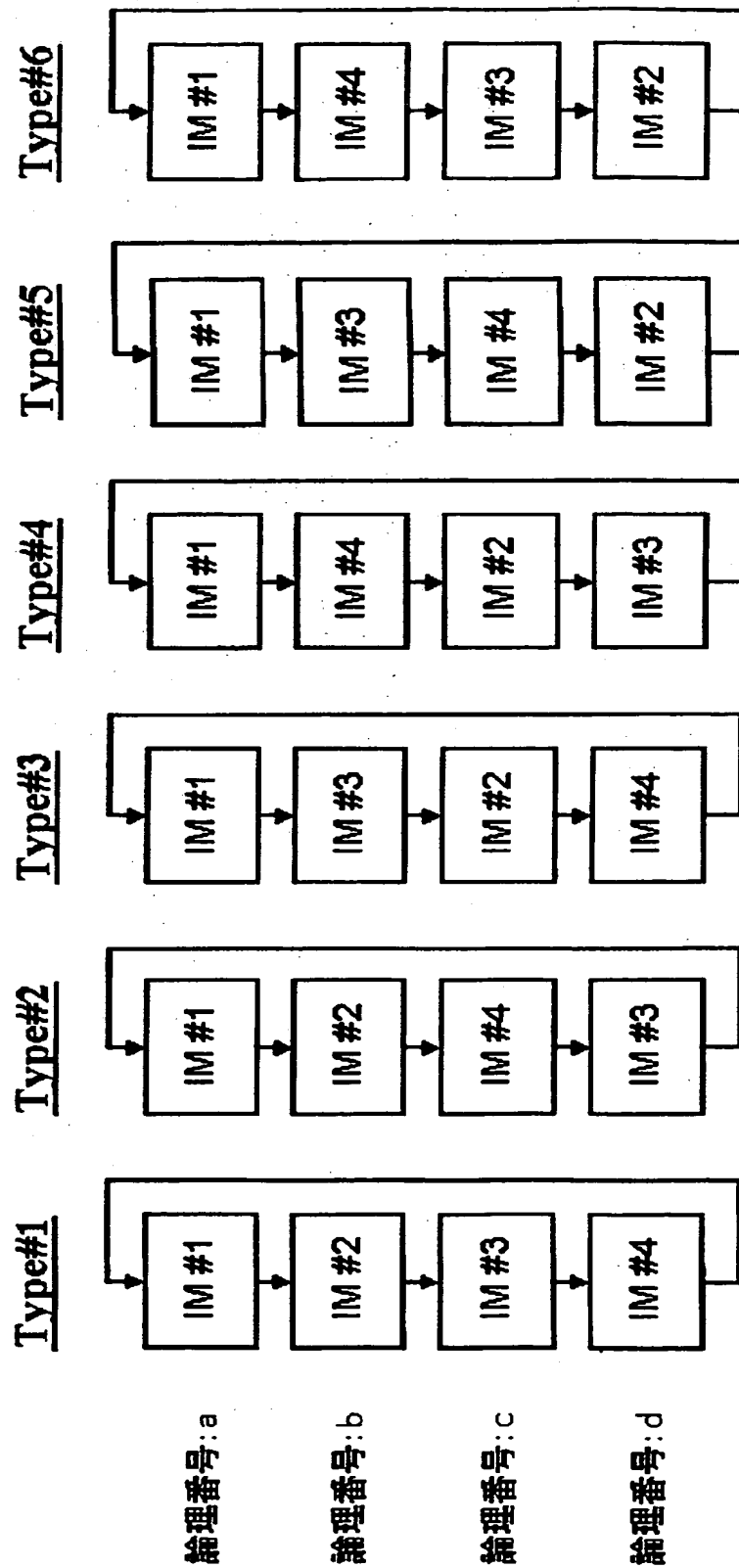
- 1 0 - 1 ~ 1 0 - 4 モジュール
- 1 5 アロケータ
- 1 6 接続許可記憶部
- 1 7 接続許可記憶制御部
- 1 8 パタン記憶部
- 1 9 パタン読出制御部
- 3 0 切替スイッチ

【書類名】 図面

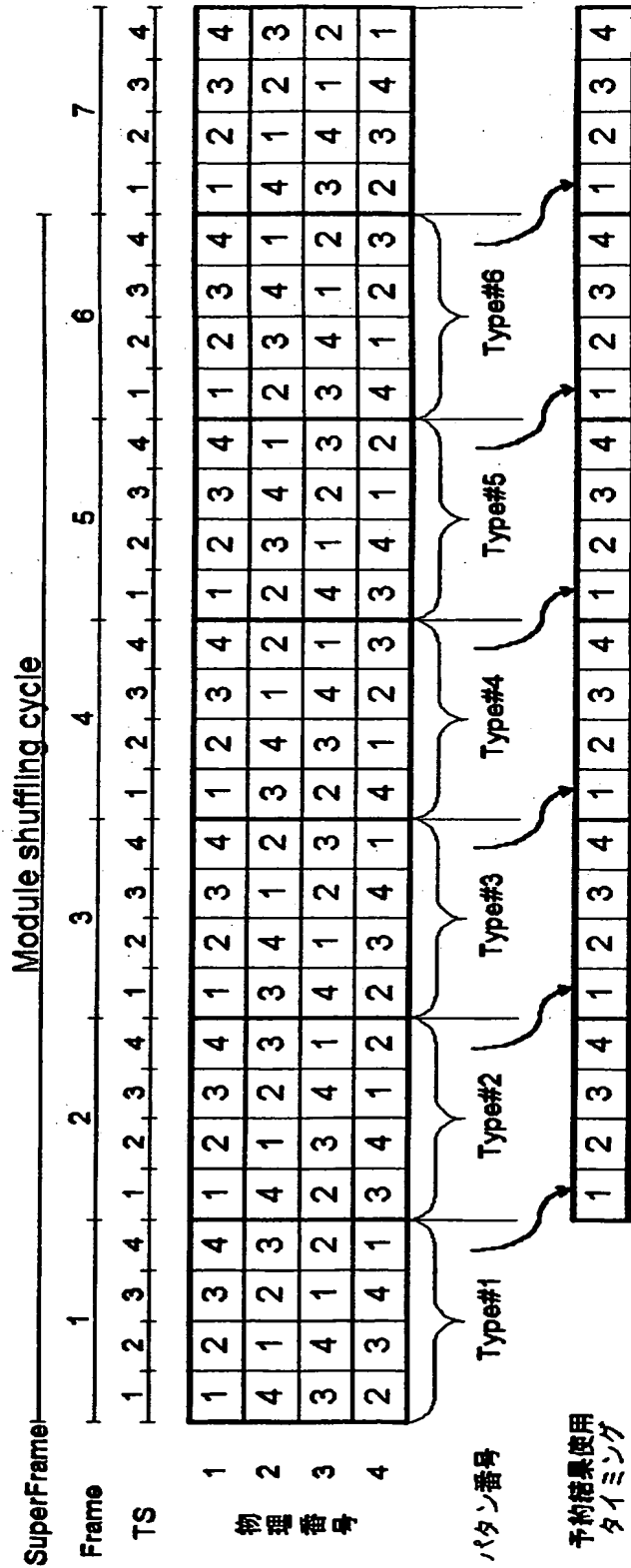
【図 1】



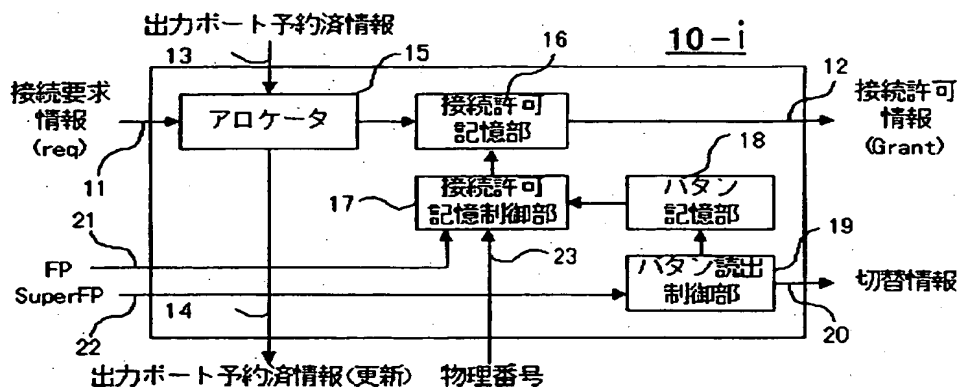
【图 2】



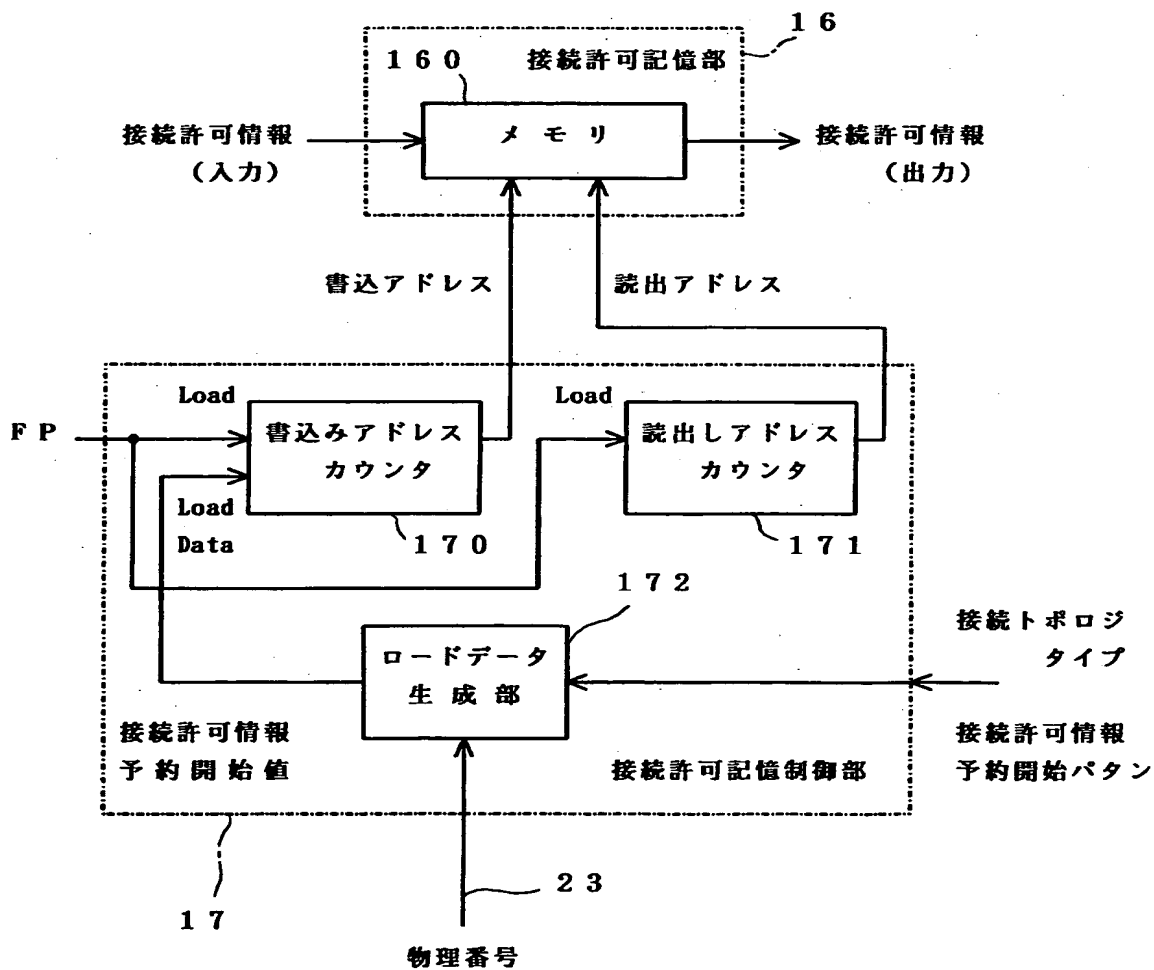
【図 3】



【図 4】



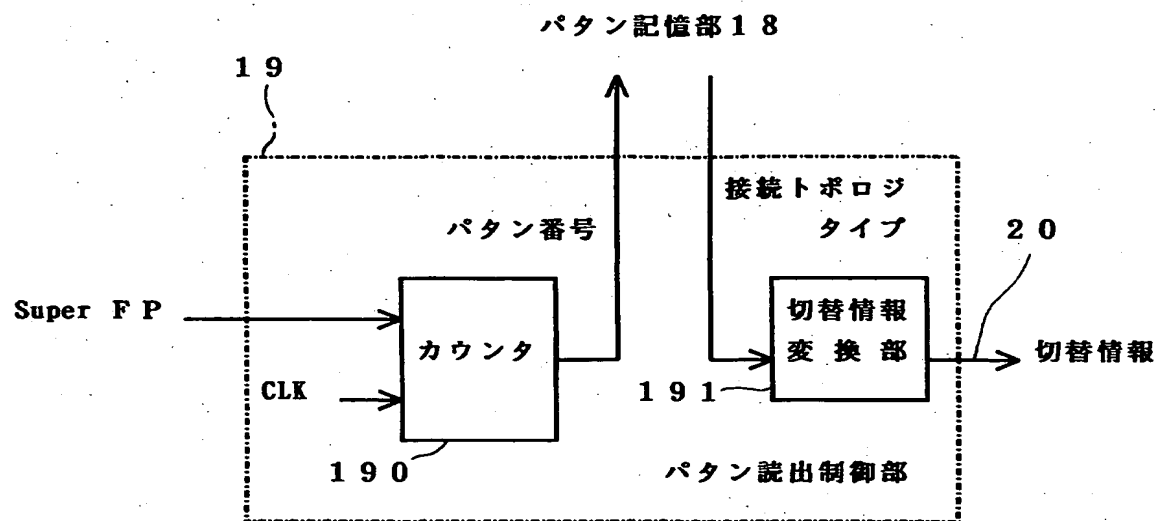
【図 5】



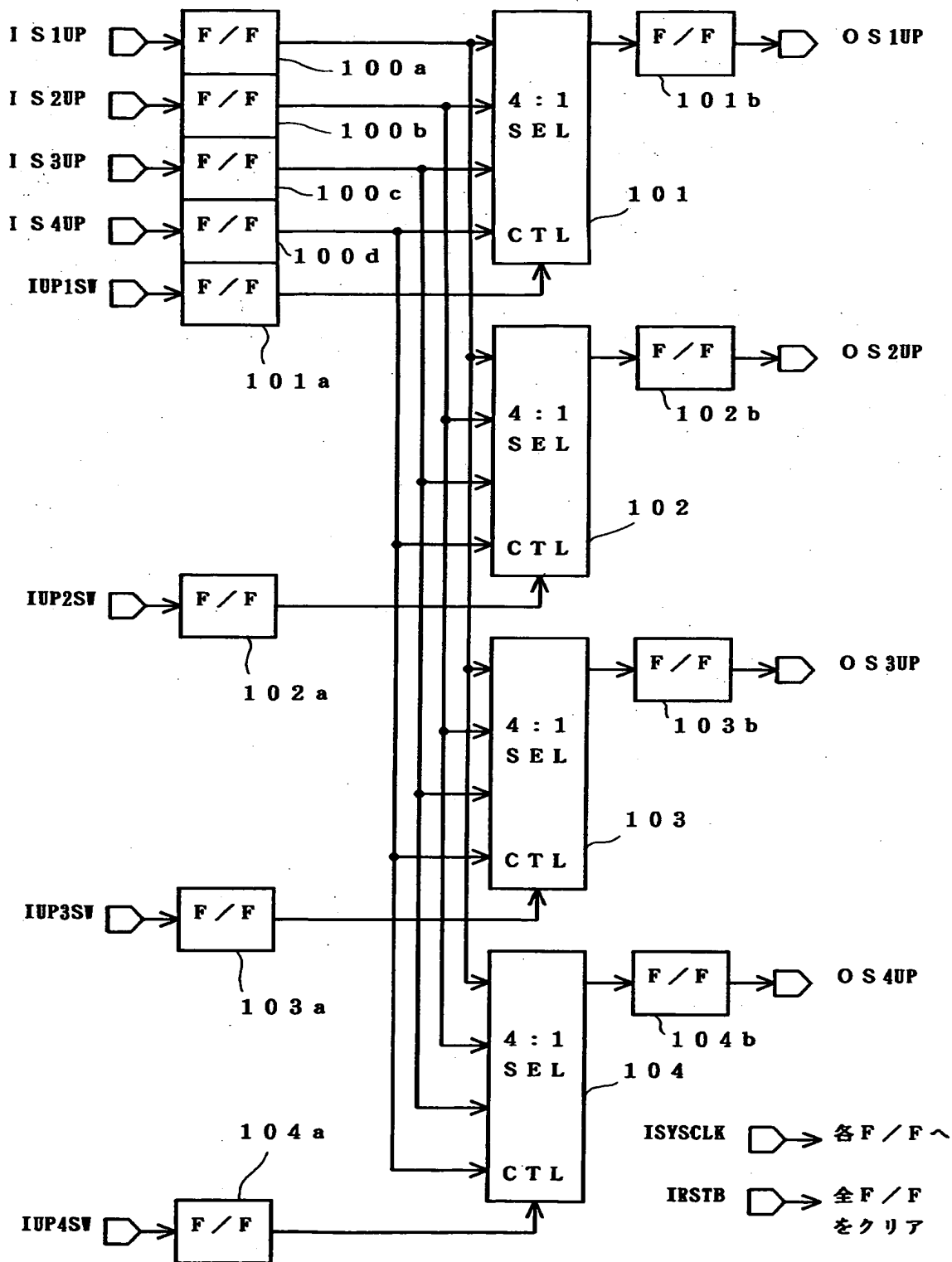
【図6】



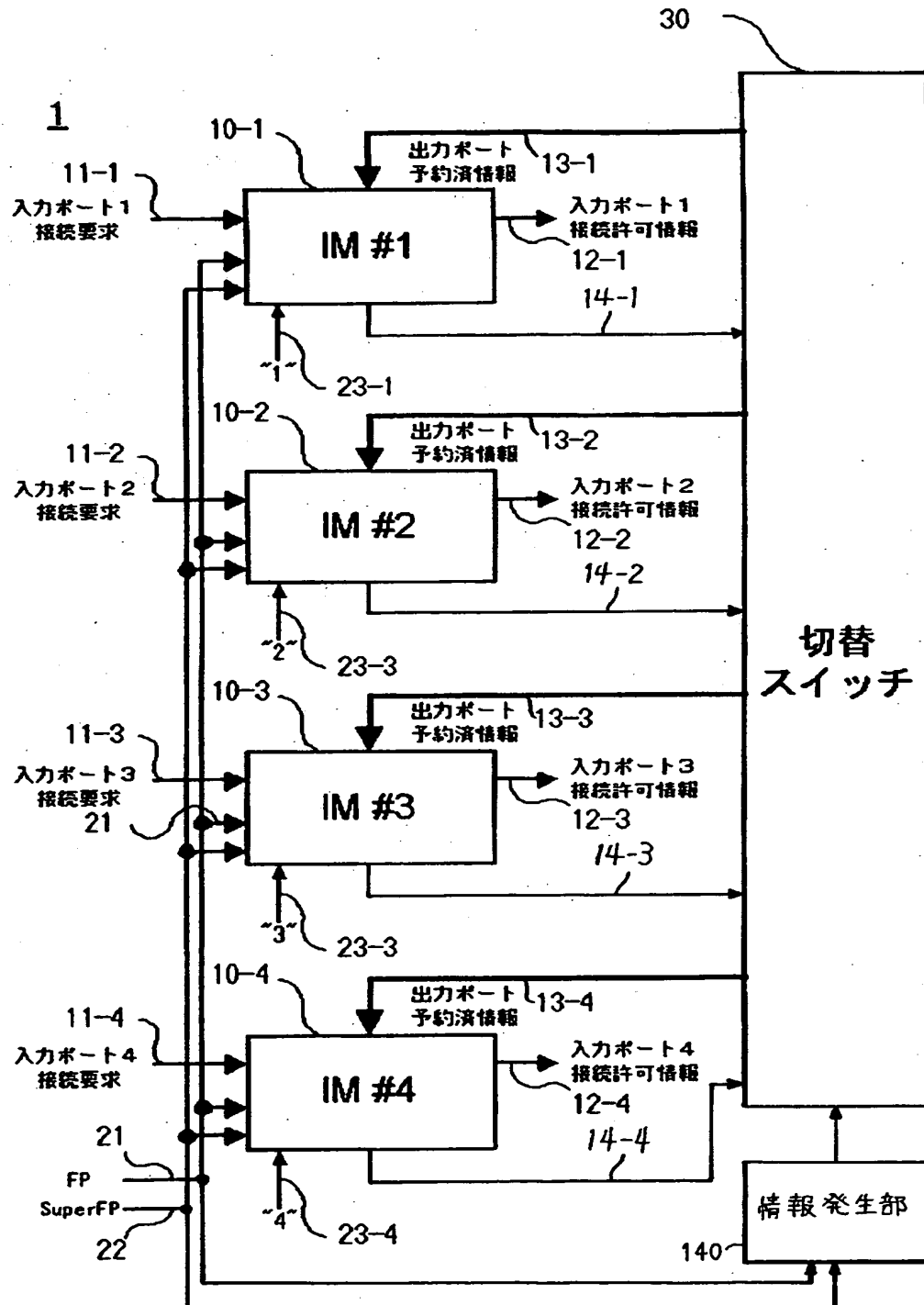
【図 7】



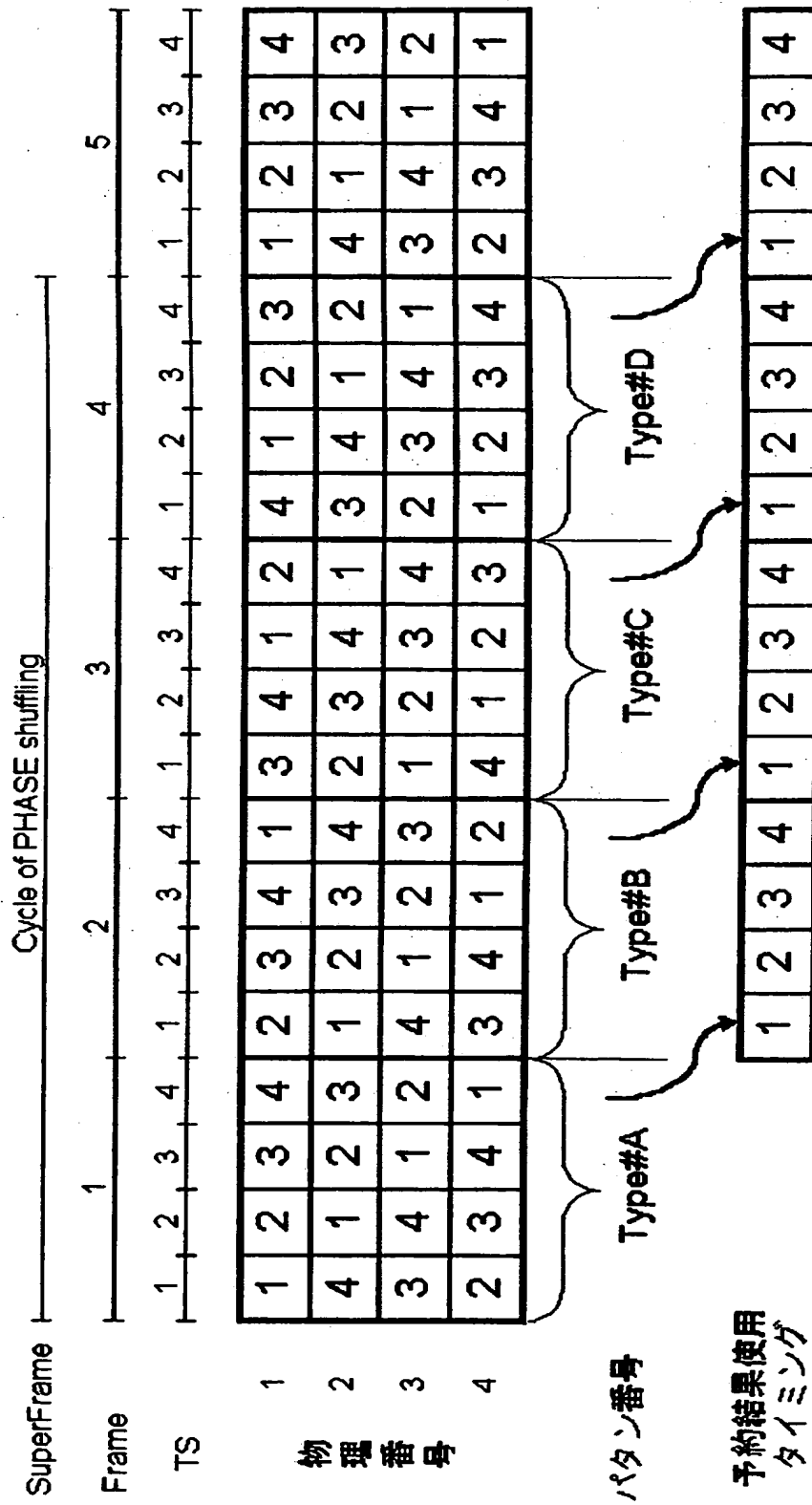
【図 8】



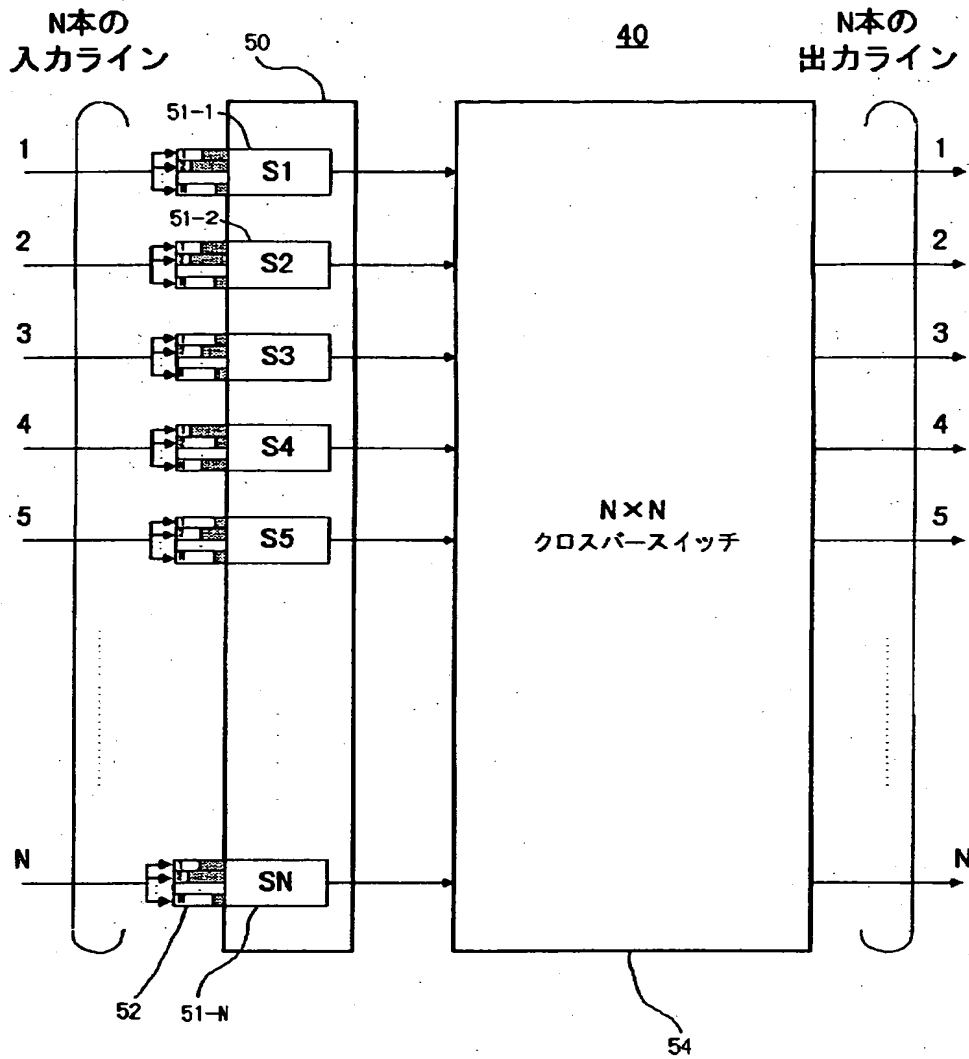
【図 9】



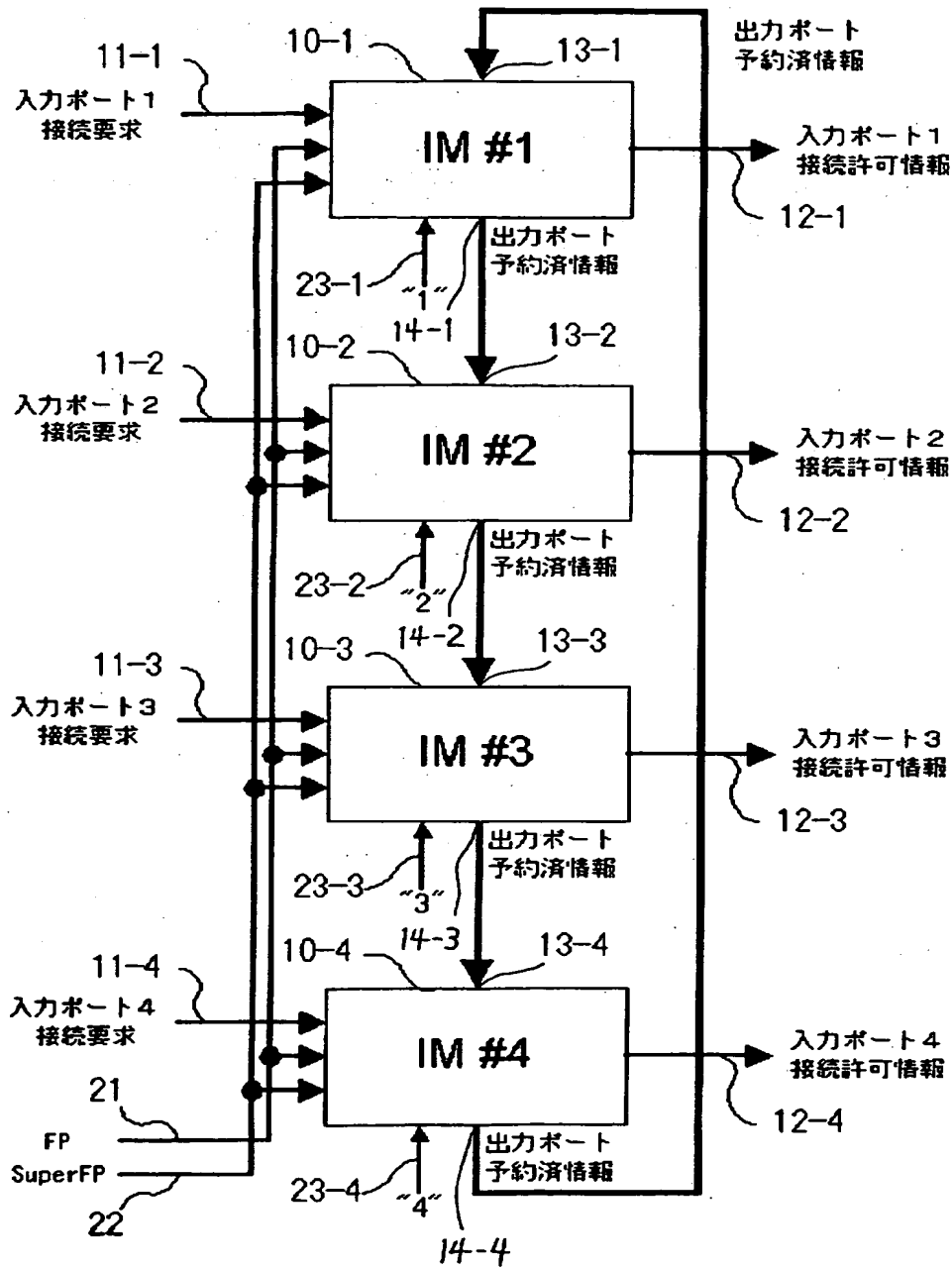
【図 10】



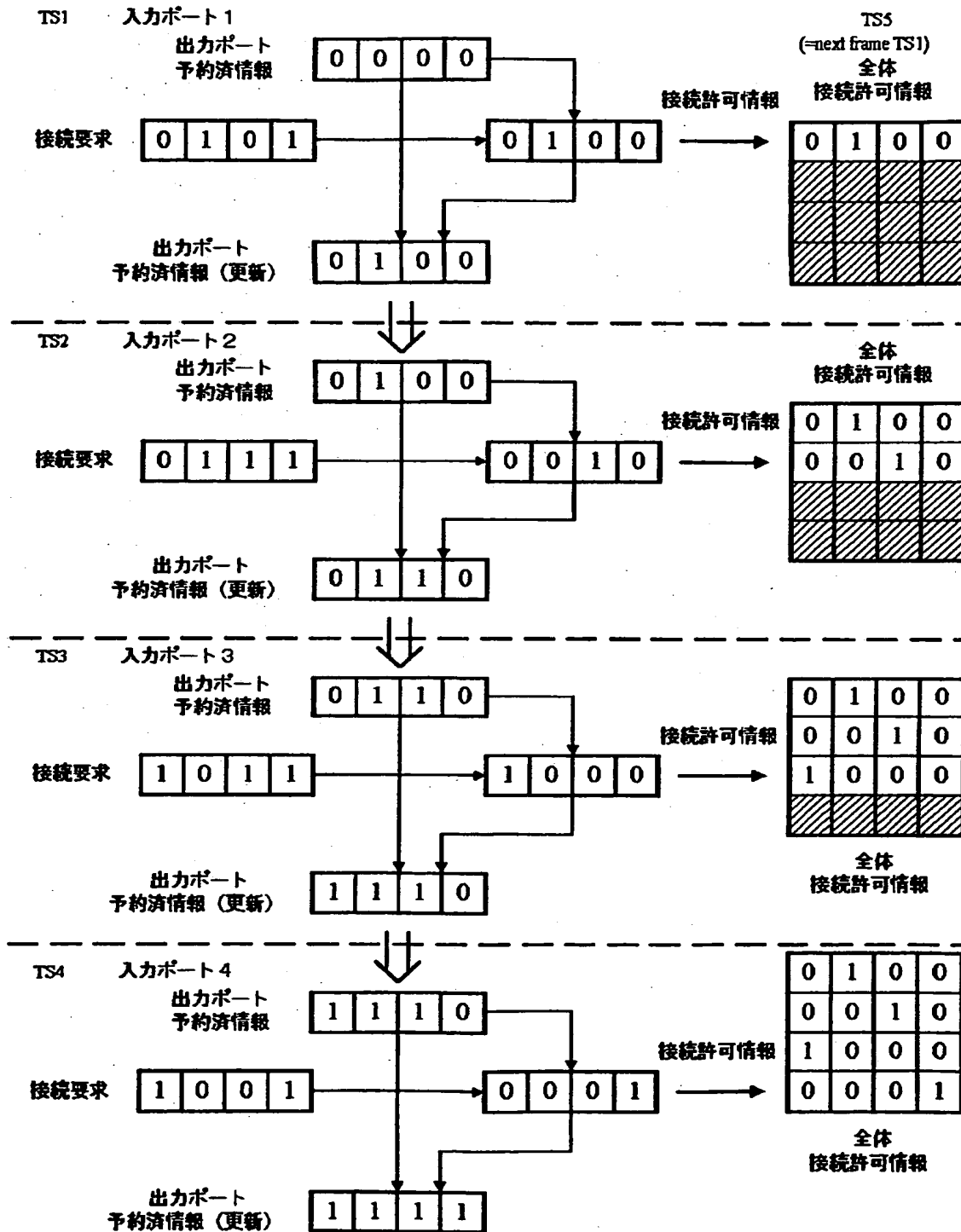
【図 1 1】



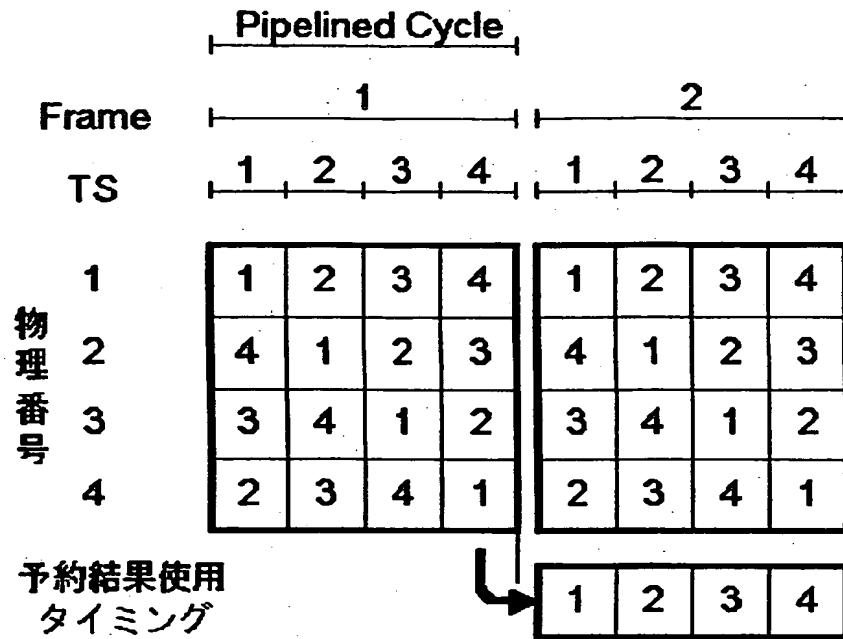
【図 1 2】



【図 1 3】



【図 1 4】



【書類名】 要約書

【要約】

【課題】 ポート間の不公平性を解消する。

【解決手段】 各モジュール 10-i ($i=1\sim 4$) の間で受け渡しをしている出力ポート予約済情報をモジュール 10-i 外の切替スイッチ 30 に入力し、出力先を変更できるようにする。モジュール 10-i と切替スイッチは、モジュール 10-i が複数タイムスロットの接続許可処理を行っているフレームを単位に同期して動作し、フレーム単位に接続トポロジを変更し全接続トポロジのパターンが出現するように動作する。

【効果】 接続トポロジが変更されることにより、様々な隣接ポートの組み合わせが出現し、固定化されていた入力ポートに関する優先度がシャッフルされることになり、入力ポートの予約機会に関する不公平性が解消される。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000004237]

1. 変更年月日 1990年 8月29日

[変更理由] 新規登録

住 所 東京都港区芝五丁目7番1号

氏 名 日本電気株式会社